How to explain the succes of the natural sciences

Kuipers, Theo A.F.

Published in:
EPRINTS-BOOK-TITLE

Publication date:
1989

HOW TO EXPLAIN THE SUCCESS OF THE NATURAL SCIENCES

Theo A. F. Kuipers
Department of Philosophy, Science Studies Unit, Groningen

Neglecting safe reformulations, it will be shown in Section 2 that and how the phenomenon that one theory is at a certain time more successful than another, i.e. empirical progress, can be explained by the comparative hypothesis that the first theory is closer to the truth than the second, to be defined in Section 1. In other words, empirical progress is explained on the basis of a truth-approximation hypothesis, where the latter can be seen as an explication of the intuitive idea of theoretical progress. Methodological rules, extrapolation to stratified theories and the generality of the analysis are the respective subjects of the Sections 3, 4 and 5.

1. The metascientific program of conceptually relative theoretical realism (CRTR).

The *basic principle* (BP) of CRTR is the following: assuming a domain D of natural phenomena (states/situations/systems) and a set M of *conceptual possibilities* designed to characterize them, there is a unique, time-independent subset X of all *empirical possibilities*, being the great unknown and hence the target of theory-directed research in the domain.

Remarks. 1. M is the *conceptual frame* of a research program for D, and is, technically speaking, a set of structures of a certain similarity type. *2.* Obviously, X is M-dependent (*conceptually relative*), but nevertheless also reality-dependent (*realism*). *3.* However, X does not represent "the actual world" (no descriptive realism), it is the target of our theorizing (*theoretical realism*). Combining Remarks 2 and 3 we might, in line of Putnam, call CRTR also theoretical *internal* realism (TIR). *4.* In the best traditions of empirical theories the present metatheory has a theoretical term X, indicating the (main) hypothetical entity of a program. *5.* BP needs extrapolation of the present deterministic presuppositions (due to the dichotomy empirically possible/impossible) to probabilistic assumptions. *6.* Extrapolation to social domains is problematic because X, as a set of social possibilities, will not be time-independent. Unfortunately, probabilistic extrapolation does not seem to be a remedy for this dependence.

BP is the core of a research program (CRTR) with many refinements, specializations and applications (see Kuipers 1982, 1984, 1987a, b, c, 1988). The explanation of the comparative success of theories is one of them, for which we will now prepare the ground by suitable explication of intuitive notions.

A *theory A* is conceived as a subset A of M together with the claim that A = X, i.e. the claim that A is the (perfect) characterization of the unknown X. Of course, we will call a theory true when its claim is true and false when its claim is false, with the immediate consequence that there is just one true theory, viz. theory X. For this reason X may well be called *the truth*.

Theory B is *at least as close to the truth as* A iff the symmetric difference between the sets B and X, which is a kind of set-theoretic distance between B and X, is a subset of the symmetric difference between sets A and X. This can be split into the *instantial clause* ($A \cap X \subseteq B \cap X$), telling that all possibilities rightly admitted by A are admitted by B, and the *explanatory clause* ($B\text{-}X \subseteq A\text{-}X$), of which the name will be motivated later, telling that all possibilities wrongly admitted by B are admitted by A.

The claim that one theory is at least as close to the truth as another, in the above sense, is a perfect example of an empirically testable comparative hypothesis, to be called the *truth-approximation (TA)-hypothesis*, for as we will see, it predicts, and hence explains, that the first will always be at least as successful as the second.

The definition of 'at least as close to the truth' can be transformed into a definition of 'closer to the truth' by requiring that the indicated subset is a *proper* subset. In this way we get the possibility of a sequence of theories converging to the truth.

2. *The CRT-realist explanation of comparative success of theories*

A number of preparations are needed. A *hypothesis H* is conceived as a weak version of a theory: H is a subset of M, and the claim is that X is a subset of H, $X \subseteq H$. Again, hypothesis H is true/false when the claim is true/false, respectively. Note that there is more than one true hypothesis.

The set of hypotheses $\mathcal{H}(A)$ following from theory A can be represented by the set of subsets of M which include A, for the associated hypothesis-claims follow from the theory-claim. Now it is not difficult to prove the 'bridge-theorem' that the explanatory clause ($B-X \subseteq A-X$) of Section 1 is equivalent to the condition that all true hypotheses following from theory A follow also from theory B (formally: $\mathcal{H}(A) \cap \mathcal{H}(X) \subseteq \mathcal{H}(B) \cap \mathcal{H}(X)$), which explains the name of the clause.

The data at a certain moment t, to be accounted for by our theories, can be represented as follows. Let R(t) indicate the set of realized possibilities at t, i.e. *the instances to be admitted*, and let $\mathcal{L}(t)$ represent the set of accepted hypotheses at t, i.e. *the laws to be explained*. On the basis of $\mathcal{L}(t)$ we define of course *the strongest law* to be explained as the hypothesis S(t) associated with the intersection of the sets constituting the accepted hypotheses. Of course, $\mathcal{L}(t)$ is in some way or other based on R(t), minimally we may assume that $R(t) \subseteq L$ for all $L \in \mathcal{L}(t)$, with the consequence that $R(t) \subseteq X \subseteq S(t)$, telling that R(t) does not contain mistakes and that hypothesis S(t) is true.

Now it is possible to explicate the success and problems of a theory A at time t (purposefully dressed in Laudanlike terminology). $A \cap R(t)$ indicates the set of realized examples, the *instantial success*, whereas $R(t)-A$ indicates the set of realized counterexamples, the *instantial problems*. $\mathcal{H}(A) \cap \mathcal{L}(t)$ represents the explained laws, the *explanatory success* and $\mathcal{L}(t) - \mathcal{H}(A)$ the unexplained laws, the *explanatory problems*.

For comparative judgements of the success of theories the following two clauses are obvious. Theory B is *instantially at least as successful as* theory A iff the instantial success of A is a subset of that of B ($A \cap R(t) \subseteq B \cap R(t)$), which is equivalent to the condition that the instantial problems of B form a subset of those of A. Theory B is *explanatorily at least as successful as* theory A iff the explanatory success of A is a subset of that of B ($\mathcal{H}(A) \cap \mathcal{L}(t) \subseteq \mathcal{H}(B) \cap \mathcal{L}(t)$), which is again equivalent to the condition that the explanatory problems of B form a subset of those of A. On the basis of a variant of the earlier mentioned bridge-theorem it is easy to prove that the explanatory success clause is equivalent to the condition that the in view of S(t) wrongly admitted possibilities by B are also admitted by A ($B-S(t) \subseteq A-S(t)$). The conjunction of the instantial and explanatory clause forms of course the general definition of the statement that one theory is at a certain time at least as successful as another.

Now we arrive at the basic argument of the paper: when one theory is at a certain moment at least as successful as another this can be derived from, and hence explained by, the following two hypotheses: the TA-(truth-approximation)hypothesis that the first is at least as close to the truth as the second, and the hypothesis that the data are sound. All notions in this argument have been explicated, and the proof of its validity is only a matter of elementary set-theoretical manipulation. It is also easy to prove in addition that, if both hypotheses are true, the first theory will always remain at least as successful as the second, i.e. also in the face of new evidence.

As a rule one theory will not include all instantial success of another and/or not all explanatory success, let alone both forms of success. The idea is of course that the relative merits can now be explained on the basis of a detailed analysis of the relative 'positions' to the truth, but for these cases a general argument is obviously not possible.

The basic argument makes clear that and how empirical progress is possible within a conceptual frame M for a domain D. It is important to note that the specific TA-hypothesis fundamentally presupposes that BP is true with respect to ⟨D, M⟩. This application of BP creates as it were the possibility that there may occur theories closer to the truth than others and that if theories are more successful than others it may be (but need not be) for that reason. In other words, although each specific example of empirical progress is explained on the basis of the corresponding specific TA-hypothesis, the generic phenomenon of empirical progress is explained on the basis of BP applied to ⟨D, M⟩.

Two successive generalizations bring us to the explanation of the success of the natural sciences in general. First, BP is true for all possible conceptual frames M with respect to the natural domain D, i.e. BP is true for D. Second, BP is true for all natural domains. Of course we do not claim that these generalizations do not have exceptions. If they are true in the majority of cases, they serve their purpose.

3. *Methodological rules*

It is not difficult to prove that the following *rule of success* (RS) "If theory B is more successful than theory A, then choose theory B (in principle)" is functional for approaching the truth in the weak sense that A cannot be closer to the truth than B. Hence, RS can be justified on the basis of BP as a prescriptive rule and can in our opinion be conceived as the explication of the hall-mark of scientific rationality.

The following rules are heuristic rules, of which it is easy to see that they stimulate new applications of RS: the *rule of content* (RC) "Aim at success preserving strengthening *or*, pace Popper, weakening of your theory", and the *rule of testing* (RT) "Aim at establishing new counterexamples of your theory, and new laws which cannot be explained by your theory". Finally, we like to mention the *rule of dialectics* (RD) for two theories that escape RS because of devided success "Aim at a success preserving synthesis of two RS-escaping theories".

It is important to note that RS is not a rule of inference in the sense that it does not conclude that the more successful theory is true (as hypothesis, let alone as theory). Putting RS in its generalized form "Choose the most successful theory among the available theories", it suggests at most the provisional conclusion that the chosen theory is the closest to the truth. This rule might be called the rule of TA-inference (as opposed to truth-inference), in comparison to the so-called "Inference to the best explanation", which prescribes to conclude, provisionally, that the most successful theory is true (as hypothesis), provided it has not yet been falsified. Several conceivable objections to and limitations of this rule of truth-inference do not apply to the suggested rule of TA-inference. In discussion with Van Fraassen we came to the conclusion that the main ones are: first, our rule is not restricted to the case that the most successful theory has not been falsified, and second, that being-the-closest-to-the-truth is always relative to other (i.e. the available) theories is self-evident, but that being-true(-as-hypothesis) is relative to other theories is close to a contradiction.

4. *Success of stratified theories*

Thusfar it might seem that our conceptually relative point of departure leads to an exotic form of relativistic realism. However, this would only be the case if we would exclude constraints between different conceptual frames for the same domain. In this section we will deal with the relation between an *observational* and a *theoretical* conceptual frame for our domain, where the distinction between observational and theoretical components is of course assumed to be not of the classical, absolute form but of a sophisticated, relativized kind.

Let M_t indicate the set of *conceptual-theoretical possibilities* and M_o the set of *conceptual-observational possibilities*, related by a projection function p from M_t onto M_o such that for $x \in M_t$ p(x) is, technically speaking, a substructure of x. Hence, members of M_t are supposed to include the observational components. For $A \subseteq M_t$, let pA indicate the set of all projections of all members of A.

Application of BP leads directly to the existence of unique, time-independent subsets $X(M_t) = X_t$ of M_t, called the set of *empirical-theoretical possibilities*, and $X(M_o) = X_o$ of M_o, called the set of *empirical-observational possibilities*. But now it is plausible to assume in addition the *principle of X-conservation*: $pX_t \subseteq X_o$. That pX_t should be a subset of X_o is just a semantic matter: if something is empirically possible, its observational part should also be empirically possible.

In the present context of stratified theories observational success is of course expressed in terms of subsets R(t) and S(t) of M_o, indicating the set of *realized empirical-observational possibilities*, and *the strongest observational law*, respectively. The explanation of the observational success has to be split into two components. That theory B is on the observational level explanatorily at least as successful as theory A (formally: $pB\text{-}S(t) \subseteq pA\text{-}S(t)$), follows from the hypothesis that "B is explanatorily at least as close to X_t as A" (formally: $B\text{-}X_t \subseteq A\text{-}X_t$). For this hypothesis implies that pB is explanatorily at least as close to pX_t as pA (formally: $pB\text{-}pX_t \subseteq pA\text{-}pX_t$), which implies, using X-conservation, that pB is explanatorily at least as close to X_o as pA (formally: $pB\text{-}X_o \subseteq pA\text{-}X_o$), which implies on its turn the desired conclusion if we assume in addition that all laws to be explained are true as hypothesis or, equivalently, that $X_o \subseteq S(t)$.

Due to the many-one character of projection the corresponding argument cannot be used for the explanation of observational-instantial success. Assuming that R(t) is sound, i.e. $R(t) \subseteq X_o$, the necessary and sufficient condition to prove that theory B is on the observational level instantially at least as successful as theory A (formally: $pA \cap R(t) \subseteq pB \cap R(t)$) is the hypothesis that pB is instantially at least as close to X_o as pA (formally: $pA \cap X \subseteq pB \cap X_o$). This hypothesis essentially says that the representational power of B with respect to X_o is at least as large as that of A: for all $x \in X_o$ if there is $a \in A$ with $p(a)= x$ then there is $b \in B$ with $p(b) = x$. Of course, more instantial success on the theoretical level remains explainable in terms of instantially closer to the truth.

5. *Generality speculations*

X-conservation is the first constraint leading away from absolute relativism of conceptual frames. In this final section we will formulate some notions that make it possible to explore still stronger non-relativistic positions.

A conceptual frame M for domain D (being the set of all structures of a certain similarity type), with empirical possibilities $X(M) \subseteq M$ due to BP, is a (*conceptual*) *level* for D if $X(M') = X(M)$ for all conceptual frames M' that include M ($M' \supseteq M$). E.g., the members of M' might contain a relation as component where the members of M have just a function that happens to be sufficient to characterize X. From now on M, M_o and M_t will indicate conceptual levels for D, where M_o

and M_t fulfil, in line with Section 4, minimally the following relation.

M_t is a *superlevel* of M_o (and M_o a *sublevel* of M_t) if there is a projection function p from M_t onto M_o such that p(x) is a substructure of $x \in M_t$ and $pX_t \subseteq X_o$, i.e. X-conservation, where X_t and X_o indicate $X(M_t)$ and $X(M_o)$, respectively.

Theory A_t ($A_t \subseteq M_t$) *reproduces* theory A_o ($A_o \subseteq M_o$) if $p^{-1}A_o =_{df} \{x \in M_t / p(x) \in A_o\} = A_t$, which implies, due to the fact that $pp^{-1} A_o = A_o$ holds in general, $pA_t = A_o$. It is not difficult to prove the following *commensurability theorem*: for any two levels there is at least one common superlevel (possibly trivially defined by 'concatenation' of tuples) on which all theories can be reproduced and, hence, can be compared.

M_t is *complete* with respect to M_o if $pX_t \supseteq X_o$ (hence, with X-conservation, $pX_t = X_o$). Note that if M_t introduces non-referring terms then it is likely to be incomplete with respect to M_o. Mt *reproduces* M_o if X_t reproduces X_o, i.e $p^{-1}X_o = X_t$, where the substantial condition in addition to X-conservation is: $p^{-1}X_o \subseteq X_t$, for X-conservation ($pX_t \subseteq X_o$) implies directly $p^{-1}X_o \supseteq X_t$. It is easy to check that if M_t reproduces M_o this implies that M_t is complete w.r.t M_o. Note also that if M_t introduces irrelevant components it will just reproduce M_o. M_t is (*empirically*) *richer than* M_o if M_t is complete w.r.t. M_o, but does not just reproduce it. Hence, $pX_t = X_o$ (implying $p^{-1}X_o \supseteq X_t$), but $p^{-1} X_o \nsubseteq X_t$, hence not $p^{-1}X_o = X_t$.

M is *unrestrictable* if it is richer than all its sublevels. M is *exhaustive* if it has no richer superlevel (and hence no unrestrictable superlevel). Note that if M an M' are both exhaustive (and not such that $M \subseteq / \supseteq M'$) then the one will be a sublevel of the other.

Finally, M is *optimal* if M is unrestrictable and exhaustive, i.e. in suggestive terms X(M) is the whole truth and nothing but the truth.

In the present approach the classical *ideal language assumption* (ILA) can be formulated as follows: for every domain there is an optimal conceptual level. This seems to be the strongest non-relativistic position. Of course, ILA presupposes BP, hence for the social sciences it is at least as problematic as BP itself. But ILA is certainly also problematic for the natural sciences. Although ILA may occasionally be fruitful as an heuristic principle, the following conflicting principle seems to be defended more frequently as a guide (in particular by Popper): although there may exist restrictable levels, there are no exhaustive levels, hence, no optimal levels, i.e. for every level one can find a richer superlevel. Let us call this methodological principle the *refinement principle* (RP). Every fortunate application of RP does not only lead to new types of empirical success and hence empirical progress but also, together with our basic principle BP, to the explanation of these phenomena.

REFERENCES

Kuipers, T., "Approaching descriptive and theoretical truth", *Erkenntnis* 18 (1982), pp. 343-387.

_____, "Approaching the truth with the rule of success", *Philosophia Naturalis* 21 (1984), pp. 244-253.

_____, "A structuralist approach to truthlikeness", in: *What is closer-to-the-truth?*, ed. Kuipers, T. (Amsterdam, 1987a), pp. 79-99.

_____, "Truthlikeness of stratified theories", in: *What is closer-to-the-truth?*, ed. Kuipers, T. (Amsterdam, 1987b), pp. 244-253.

_____, "Truthlikeness and the correspondence theory of truth", in: *Logic, Philosophy of Science and Epistemology* (Proc. 11th Wittgenstein-Symposium); ed. Weingartner, P./Schurz, G. (Vienna, 1987c), pp. 171-176.

_____, "Structuralistische explicaties van dialectische begrippen", in: *Congresbundel 9e Filosofiedag* 1987 (Delft, 1988), pp. 191-197.

* * *